

**Discipline:** Methods

**1. Language**

English

**2. Title**

Applied Regression Analysis

**3. Lecturer**

Prof. Georg von Graevenitz, Ph.D., Faculty, (Queen Mary University of London),  
<https://www.qmul.ac.uk/busman/staff/academic/profiles/vongraevenitzg.html>

Prof. Dr. Stefan Wagner, Faculty, (ESMT Berlin),  
<http://www.esmt.org/eng/faculty-research/stefan-wagner/stefan.wagner@esmt.org>

Onsite support: tba, PhD candidate, ESMT

**4. Date and Location**

ESMT Berlin

Room: tba

Schlossplatz 1, 10178 Berlin

<https://www.esmt.org/school-networks/about-esmt/campus/esmt-campus-berlin>

11.6. – 14.6.2024

**5. Course Description**

5.1 Abstract and Learning Objectives

This course will provide participants who have basic skills in statistics and econometrics with an introduction to current core methods used in the analysis of observational, experimental, and quasi-experimental data. The methods covered are widely used in economics and are increasingly also required for good publications in top management journals.

The aim is to cover theory and selected applications, but more importantly to introduce participants to the use of statistical software that will allow them to apply the methods discussed in the course to data.

This course covers important methods used in the multivariate analysis of data. The course revisits basic concepts of the linear regression model and its properties and covers selected advanced topics such as the analysis of duration data and (quasi) experimental designs as well as methods to deal with the problems of endogeneity and sample selection. The theoretical basis of these methods is discussed but the focus of the course is on the application of the methods to data sets. Applications will be studied with the help of data provided by the lecturers and with reference to recent publications.

After the course, participants will...

- have a basic understanding of the theoretical underpinnings of multiple regression models.
- be able to apply regression methods to the investigation of economic relationships and processes.
- understand the econometric methods, approaches, ideas, results and conclusions met in the majority of economic books and articles.
- be aware of common pitfalls and mistakes to avoid when conducting regression analysis.
- be able to use the software program STATA to carry out empirical analysis based on regression analysis.

## 5.2 Content

The course covers several important methods and approaches to econometric analysis. The derivation and proofs of basic formulas and models are presented which allows students to understand principles of econometric theory. The main emphasis of the course is on the economic interpretations and applications of considered econometric models. The methods selected are increasingly used in applied research by management scholars. In addition to standard cross-section models, we cover selection and duration models as well as experimental designs in econometrics.

Participants will be actively involved with computer exercises in this course using the STATA econometrics program. Throughout the course, they will use STATA to implement a series of econometrics exercises designed to provide experience with various tests and estimation procedures. The instructions for these exercises and the data required for their implementation will be sent to participants weeks before the course.

## 5.3 Schedule (including start and end time)

### Day I (11.6.2024)

- 9:00 – 9:30 Arrival of participants, reception, check-in, and introduction
- 9:30 – 11:00 Introduction and overview – 1st steps with STATA
- 11:30 – 13:00 Recap: Basic concepts from statistics (random variable, statistical inference, point estimators, and statistical simulation)
- 14:00 – 15:30 Recap: multiple regression analysis, the OLS-estimator and interpretation of coefficients

16:00 – 17:45 Exact and asymptotic properties of the OLS estimators, statistical inference in regression models

### **Day II (12.6.2024)**

- 9:30 – 11:00 Functional form, qualitative information and dummy-variable approach, and transformations in multivariate analyses
- 11:30 – 13:00 Diversions from the classical model: heteroscedasticity, omitted variables and endogeneity
- 14:00 – 15:30 Review Session I: Endogeneity & Instrumental Variables (Angrist & Lavy 1999)
- 16:00 – 17:30 Exercise session I: Probability models: Probit and Logit (Bloom et al. 2015)

### **Day III (13.6.2024)**

- 9:30 – 11:00 Models for survival times – accelerated failure time and hazard rates
- 11:30 – 13:00 Exercise session II: The Duration of Patent Examination at the European Patent Office (Harhoff/Wagner 2009).
- 14:00 – 15:30 Difference-in-Differences & Panel Data Estimators
- 16:00 – 17:30 Exercise session III: Natural- Experiments (Bloom et al. 2015)

### **Day IV (14.6.2024)**

- 8:30 – 10:00 Econometrics and machine learning
- 10:30 – 12:00 Exercise session IV: Using Lasso & Elastic-Net (Belloni et al. 2014 )
- 13:00 – 13:30 Wrap-up & Feedback
- 13:45 – 15:15 In-call exam

## 5.4 Course format

The course will consist of a combination of lectures, exercise sessions, and a final exam. Lecturers will use recent journal articles as well as book chapters to teach the participants common regression methods. Participants are advised to carefully read the assigned materials before the class.

## 6. Preparation and Literature

### 6.1 Prerequisites

The course requires basic skills in statistics and multivariate data analysis techniques. Concepts such as mean values, standard deviations and covariance matrices should be familiar to the participants. In addition, a basic understanding of regression analysis and testing procedures is helpful but not an essential requirement for understanding the contents.

The amount of materials covered during the four days is large. We primarily target participants with first own experience in data analysis. Moreover, basic knowledge of the statistical software package used in the course, STATA, will facilitate active participation.

Participants should be prepared to use their laptops during the course to directly follow important parts of the course related to the application of important concepts to data sets using the statistical software STATA. Therefore, an installation of STATA version 15.0 or higher is required. In Germany, student versions of STATA can be obtained from DPC (<http://www.dpc-software.de>) for about 100 EUR.

## 6.2 Essential Reading Material

Angrist, J. and Lavy, V. (1999): Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement, *The Quarterly Journal of Economics*, Vol. 114(2), 533-575.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "High-Dimensional Methods and Inference on Structural and Treatment Effects." *Journal of Economic Perspectives*, 28 (2): 29-50.

Bloom, N./ Lian, J./ Roberts, J. and Ying, Z. J. (2015): Does working from home work? Evidence from a Chinese experiment, *Quarterly Journal of Economics*, Vol. 130 (1), 165-218.

Chenhall, R. H./ F. Moers (2007): The Issue of Endogeneity within Theory-Based, Quantitative Management Accounting Research, *European Accounting Review*, 16, 173–196.

Harhoff, D./ Wagner, S. (2009): The Duration of Patent Examination at the European Patent Office, *Management Science*, Vol. 55 (12), 1969-1984.

Imbens, G. / Wooldridge, J. (2009): Recent Developments in the Econometrics of Program Evaluation, *Journal of Economic Literature*, 47, 5-86.

Imbens, Guido W. 2020. "Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics." *Journal of Economic Literature*, 58 (4): 1129-79.

We also ask the participants to familiarize themselves with STATA before the course starts. There are excellent tutorials available online at

<http://www.stata.com/links/resources-for-learning-stata/>

<http://www.ats.ucla.edu/stat/stata/>

[https://scholar.harvard.edu/files/mcgovern/files/practical\\_introduction\\_to\\_stata.pdf](https://scholar.harvard.edu/files/mcgovern/files/practical_introduction_to_stata.pdf)

## 6.3 Additional Reading Material

Gujarati, D. (2008). Basic Econometrics, Mcgraw-Hill Higher Education, 4th ed.

Kennedy P. (2003). A Guide to Econometrics. MIT Press, 5th edition.

Angrist, J. D. and J. S. Pischke (2009): Mostly Harmless Econometrics: An Empiricist's Companion, Princeton Univ Pr.

Angrist, J. D., & Pischke, J. S. (2014). Mastering'metrics: the path from cause to effect. Princeton University Press.

Kiefer, N. (1988): Economic Duration Data and Hazard Functions, *Journal of Economic Literature*, Vol. 26(2), 646-679.

Stock, J./ Watson M. (2004): Introduction to Econometrics, Addison Wesley, 2nd edition.

Wooldridge, J. (2009). Introductory Econometrics: A Modern Approach. South Western College Publishing.

Wooldridge, J. (2002): Econometric Analysis of Cross Section and Panel Data, MIT Press.

If applicable: Further literature will be sent to the participants weeks prior the starting date of the course.

#### 6.4 To prepare

All participants are required to read the essential reading material before the course. We will rely largely on the paper by Bloom et al. (2015) to illustrate various tools of multivariate data analysis and we will provide some of the data used in this paper to the participants for the exercise sessions. For this reason, we ask that all participants familiarize themselves with the paper **and prepare a short summary and discussion (appr. 5 minutes, 3 slides) of Bloom et al. (2015)**. We will invite one or two volunteers to present their thoughts to the group.

### 7. Administration

#### 7.1 Max. number of participants

The number of participants is limited to 20.

#### 7.2 Exam

A 90-minute in-class exam will be offered at day IV.

#### 7.3 Credits

The course (including the exam) is eligible for 6 ECTS.

### 8. Arbeitszeitaufwand / Working Hours

Working Hours	hours
<i>Active participation</i>	32 hours
<i>Preparations</i>	74 hours
<i>Follow-up individual study period</i>	74 hours
<b>SUMME</b>	<b>180 h</b>
<b>ECTS: 6</b>	